Assessing the Status of Geologic Map Coverage of the United States—A New Application of the National Geologic Map Database

By David R. Soller

U.S. Geological Survey
926-A National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6977
e-mail: drsoller@usgs.gov

INTRODUCTION

Systematic geologic mapping of the United States has been conducted for more than 125 years. In the period centered on 1895-1920, the USGS conducted the first such program, the Geologic Atlas of the United States. The Atlas included about 230 products, at scales ranging from 1:14,400 to 1:250,000 (Figure 1). It is notable that the scientific and cartographic standards developed to guide that mapping (Powell, 1888) have, with modest revision, endured to this day.

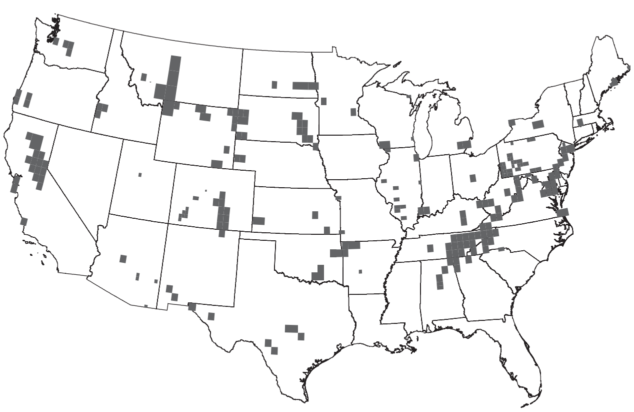


Figure 1. Index map of the USGS series "Geologic Atlas of the United States" (ca. 1895-1920). Information derived from the NGMDB Map Catalog.

In more recent times, geologic mapping has been conducted under many programs in Federal, State, and other agencies. These programs have differed in emphasis owing to funding source and time period; partly as the result, geologic maps vary significantly in content and format. These differences present a real challenge to the preparation of index maps that purport to show geologic map coverage – if all geologic maps are not alike in content, scale, detail, vintage, or currency, which then should be included in an index map? By what criteria should we differentiate or classify geologic maps for this purpose?

Purpose

The principal and most obvious purpose for index maps is to convey to any user, whether a practicing geologist or a homeowner, the availability of published geologic maps. Also, the Nation's geological surveys need to know the areas for which geologic maps have been made, and when those maps were published. Such information helps each agency to prioritize areas that should be mapped in the future (or remapped, usually in more detail). Furthermore, it demonstrates to legislators and oversight agencies such as the U.S. Office of Management and Budget (OMB) that funding has produced tangible results.

In response to requests from the Association of American State Geologists (AASG) and the USGS National Cooperative Geologic Mapping Program (NCGMP), the National Geologic Map Database project (NGMDB) in 2005 developed the capability to generate index maps showing geologic map coverage at different scales and for various time periods. The information supplied by the NGMDB (i.e., index maps and numeric summaries of the extent of intermediate and large-scale geologic mapping in the U.S.) was needed in order to fulfill two OMB requirements for NCGMP "performance metrics". These metrics served as partial documentation of AASG and USGS performance in addressing the goals of the Geologic Mapping Act (http://ncgmp.usgs.gov/ncgmpabout/ngmact/). It is anticipated that such information will be required annually.

Ideally, these index maps would be created from a database designed specifically for the task. However, no such database existed. For the AASG and NCGMP, therefore, the logical choice was to extract information from the NGMDB Geoscience Map Catalog (available at http://ngmdb.usgs.gov/). The Catalog is a general-purpose database containing bibliographic records on more than 75,000 geologic maps and other geoscience maps and reports for the U.S., published by more than 350 organizations. Based on new AASG and NCGMP requirements for this information, and on the general informative value of index maps, the NGMDB project is now endeavoring to revise its database to accommodate agency requests for such information, as described below.

METHOD

The request from AASG and NCGMP was to provide index maps showing the location of “modern”, general-purpose geologic maps of intermediate and detailed scale. As noted in the Introduction, geologic maps are not all alike in content, scale, detail, vintage, or currency. Which maps should be considered modern, which are general purpose, and what is an intermediate or detailed scale? It was decided for this purpose that “modern” maps would be somewhat arbitrarily defined as those published since 1959. Because the NCGMP was created to address the goals of the Geologic Mapping Act of 1992, a secondary objective was to identify maps produced since that date. General-purpose geologic maps are those that include all geologic units in the map area and that focus on geologic history and the characteristics of the materials (for example, the typical geologic map of a quadrangle or county). The decision regarding map scales appropriate to portray was a difficult one, and so for the initial set of index maps, all maps of scale 1:250,000 and more detailed were included. Upon inspection of these index maps it became clear that the scale should be more restricted, to intermediate-scale (defined for this project as 1:100,000) and more detailed maps. Decisions such as this, necessary but arbitrary in nature, resulted in the omission of numerous useful geologic maps (e.g., county maps at scales approximately 1:125,000).

Step 1 – Selecting records to evaluate

With these definitions providing a constraint, the process of creating the index maps began with a query of the Map Catalog’s “theme” field, in order to identify all bibliographic records that contain bedrock or surficial geologic information. The problem with this approach was immediately evident – although each of these products contain geologic map information, many could not legitimately be described as general-purpose geologic maps. This problem was unavoidable because no field more relevant than “theme” was available.

Each publication in the Map Catalog is assigned one or more geologic themes that describe its content (see Figure 2 and http://ngmdb.usgs.gov/ngmdb/define.html). For example, consider a publication that addresses landslide hazards – the landslide hazard potential, or the surveyed landslides, commonly are shown on a geologic map in order to provide context for these features. The geologic map may be newly-developed by the landslide-mapping project, or it may have been reproduced in full or reduced detail from a map originally released in another publication. In the Map Catalog database, the landslide map would be assigned the geologic themes “Landslides” and either “Bedrock Geology” or “Surficial Geology” because it contains a geologic map. The purpose of the “themes” field is to assist the user in finding the type(s) of maps they need without omitting from the database search any publications that could be useful. When this database was under development (ca. 1995) this was recognized as an important feature – it returns to the user a list of all maps that might possibly be useful, thereby giving the user the opportunity to choose from a larger set of products than would be possible if only the principal theme of the product were recorded in the database.

The search for “Theme=bedrock or surficial geology” yielded 28,100 publications. Bibliographic information and bounding coordinates for each publication were exported from the Map Catalog’s Oracle database to a .DBF (v. 4) file. To simplify the handling of these records,

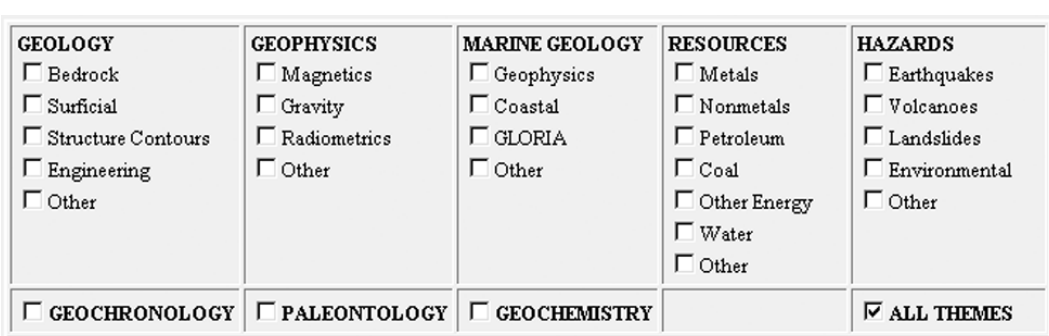


Figure 2. Geoscience themes in the Map Catalog.

the publications were divided into four files, by scale range (roughly, 1:24,000 and more detailed; 1:25,000 – 1: 99,000; 1:100,000; and 1:101,00 – 1:251,000).

Step 2 – Creating the "footprint" of each map

To simplify the entry and management of bibliographic records, and to minimize data-entry errors, the Map Catalog database contains the bounding coordinates (actually, the NW and SE corners) of each map publication rather than the actual extent of the mapped area. For maps that fill a rectilinear area (e.g., a quadrangle), the bounding coordinates accurately represent the mapped area. However, for a map of an irregular area (e.g., a "strip map" of the geology along the Chesapeake and Ohio Canal in D.C., WV, VA, and MD), the bounding coordinates can drastically over-depict the area mapped (Figure 3). Therefore, it was decided that maps of such irregular areas should not be shown by the index maps; differentiating these maps from more rectilinear maps proved difficult, as explained in Step 5, below.

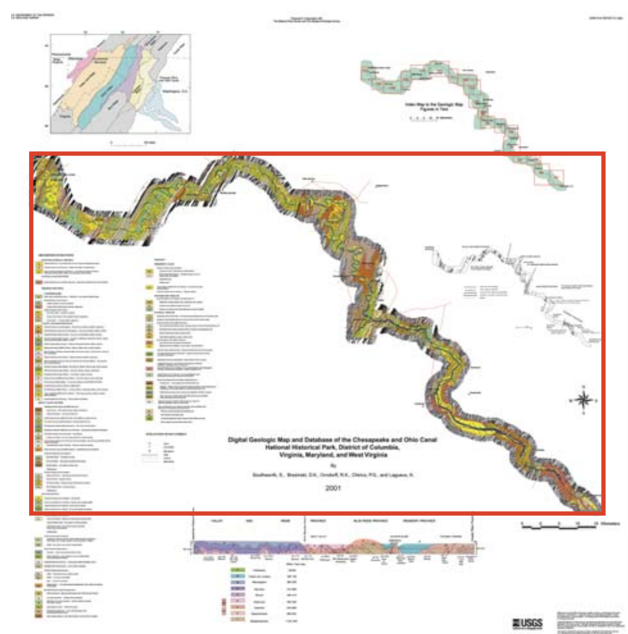
To prepare the selected records for evaluation and display, each .DBF file was imported into ArcView, using an Avenue script written by Chris Garrity (USGS) that converted each bounding coordinate pair to a polygon. Each of the four ArcView shapefiles were visually evaluated for obvious errors in the data-entry of bounding coordinates or map scale (e.g., a map of a large area that, while specified as 1:100,000, is obviously of a smaller scale, such as 1:1,000,000). This visual check was found to be a valuable supplement to the automated, logical checks performed on each newly-added Catalog record, and will therefore be included in database error-checking procedures. Erroneous coordinates and map scales were corrected, and the file then was reimported to ArcView, to be rechecked until no errors were visually detected.

Step 3 – Exploring the data

As may be apparent from Steps 1 and 2, index maps could not simply be generated by selecting the map theme "bedrock geology" or "surficial geology" and then displaying the outline of each map's bounding box. Instead, each publication had to be evaluated for suitability. This evaluation focused on the most informative field for this purpose – publication title. From the title, I hoped to differentiate:

1. general-purpose versus more specialized geologic maps (e.g., "Revision of Middle Proterozoic Yellowjacket Formation, Central Idaho"), and
2. maps of quadrangles and counties (and parishes and boroughs) versus maps of more irregular outline.

A



B

Figure 3. Comparison of the area that was geologically mapped versus the map area as described by its bounding coordinates.
A) a typical quadrangle map (Kepferle, 1973); here, the mapped area and the bounding box are the same, and so the box is a valid estimate of the mapped area.
B) a geologic map of an irregular area (Southworth et al., 2001); the bounding box (thick, dark line) is not a valid estimate of the mapped area.

Before attempting to evaluate the titles, several days were allotted to peruse the data, seemingly at random, in order to become more familiar with the content. The importance of this can scarcely be over-emphasized – if I had immediately "waded into" the evaluation process, relying on my assumptions about how to efficiently proceed, the result would have been far less accurate in portraying the status of geologic map coverage.

Step 4 – Defining the evaluation criteria

From Step 3 it was verified that useful information regarding map content and shape (i.e., general-purpose versus specialized geologic map; quadrangle versus irregular shape) could be interpreted from the publication titles. Because titles vary greatly in format and information content, and because any such interpretation inherently carries some degree of uncertainty, this process represents a necessary compromise dictated by the reality that time and effort could not be expended in order for each publication to be reviewed by the person or agency most familiar with its content.

Two database fields were added to the Map Catalog: "mapshape" and "include"; the latter considered map shape and, further, indicated whether a map would be included in the index maps. The mapshape field included these values:

- 4 = quadrangle
- 3 = county, parish, or borough
- 2 = irregular (rarely used; evaluated by finding and then inspecting the map)
- 1 = probably irregular (evaluated based on the publication title)
- 0 = not evaluated, because not a geologic map.

For the "include" field, titles pertaining to general-purpose maps of quadrangles and counties were identified and assigned to be included on the index maps. Specialized geologic maps and general-purpose geologic maps of irregular areas were assigned to be not included, with an exception: maps of islands, large coastal areas, and major parks, where quadrangle and county maps were unavailable, were assigned a value that enabled them to be considered for inclusion on the index maps. As noted in the concluding section of this paper, because selected irregularly-shaped maps were included, the area that has been geologically mapped was somewhat overestimated; this error will be minimized in future versions of these index maps.

The "include" field allowed these values:

- 4 = general-purpose geologic map of a quadrangle or county
- 3 = general-purpose geologic map of an irregular area such as an island, a coastal region, or a major park
- 2 = general-purpose geologic map of an irregular area
- 1 = specialized geologic map, or not a geologic map but containing geologic content
- 0 = does not contain geologic map information.

Examples of maps assigned to these two categories are provided in Figure 4.

Step 5 – Evaluating each map

The most efficient method for identifying appropriate maps was found to be a systematic query of map titles, in ArcView; the query searched for keywords associated with general-purpose geologic quadrangle maps or more specialized geologic maps. For example, a query might search for titles containing text strings such as "geologic map" and "quadrangle", and include these maps. Conversely, a query might search for terms that commonly are applied to specialized maps or those of irregular outline, such as "formation" or "range" or "district", and omit these. All maps selected by a query then were assigned values for the two new fields described above.

To a significant degree this systematic and logical approach identified the appropriate maps, and it greatly expedited the assignment of this new information to each publication. However, through typical error-checking procedures it became clear that this approach caused to be omitted many maps that should have been included (e.g., those that used the term "sheet" or "folio" rather than "quadrangle", or those whose quadrangle name included an omitted term such as "range"). Conversely, the process caused to be included many maps that should have been omitted (e.g., a map with a title something like "Geologic map of gold-bearing rocks within the XYZ quadrangle"). Because I could not be assured that all titles had been correctly interpreted by these queries, each title then was inspected in order to identify publications that likely had been assigned to the incorrect "include" or "mapshape" category. Corrections were made, and the ArcView files were then ready for preparation of index maps. Of the 28,100 publications originally selected, 15,026 were identified according to the criteria listed above. Because it was later determined that we would show only the maps of scale 1:100,000 and larger (i.e., more detailed), the number of relevant publications was further reduced, to 13,597.

Step 6 – The index maps

Following discussions with the AASG and NCGMP, maps classified with an "include" value of 3 or 4 were

A

Include?		Title
	2	Geology and ore deposits of the Afterthought mine, Shasta County, California
	4	Geology and mineral resources of Jones Ranch School quadrangle, McKinley County, New Mexico
	4	Geology and coal resources of Vanderwagen quadrangle, McKinley County, New Mexico
	4	Geologic map of the Lake Helen quadrangle, Big Horn and Johnson counties, Wyoming
	4	Preliminary geologic map of the Lillis Ranch quadrangle, California
NO	2	Geology of the Tejon Hills area - Arvin and Tejon Hills quadrangles, Kern County, California
NO	1	Stratigraphy, structure, and mineral deposits in the Oro Grande series near Victorville, California
YES	4	Bedrock geologic map of the Woodford quadrangle, Bennington and Windham Counties, Vermont
YES	3	Bedrock geologic map of Yosemite Valley, Yosemite National Park, California
	4	Geologic map of the Vidal, California, and Parker SW, California-Arizona quadrangles
	4	Geologic map of the Vidal NW, Vidal Junction, and parts of the Savahia Peak SW and Savahia Peak

B C

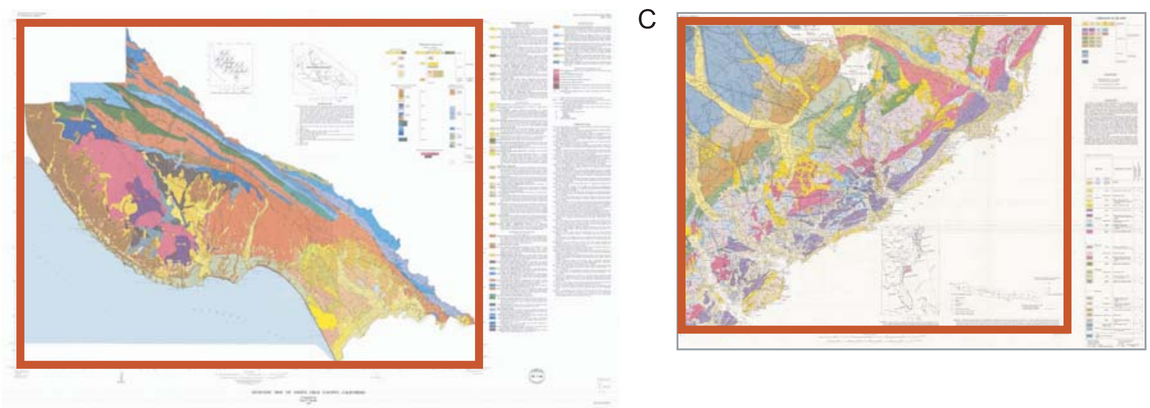


Figure 4. Examples of products included on the index maps.
A) inferring the type of geologic map, based on title ("yes" = included on index maps; refer to values for the "include" field in Step 4). Example of a general-purpose quadrangle geologic map is shown in Figure 3A.
B) general-purpose county geologic map (Brabb, 1989). Bounding box shown by thick, dark line.
C) general-purpose geologic map of irregular area (McCartan et al., 1984). Bounding box shown by thick, dark line. See Figure 3B for example of irregular-area geologic map that was not included.

incorporated into a variety of index maps. Clearly, many permutations of map scale and vintage can be displayed in such maps; three examples are provided here (Figure 5). Commentary on the extent and distribution of geologic map coverage across the Nation, although feasible and desirable, is not within the scope of this methods-oriented paper. Such commentary, if it were to be provided at some later date, would most appropriately come from the NCGMP and AASG.

CONCLUSIONS AND RECOMMENDATIONS

The NGMDB project has long recognized the need to produce, for the practicing geologist, decisionmaker, and general public, a set of accurate and detailed index maps of geologic map coverage. The requirement for performance metrics provided the impetus to focus on this need, thereby giving it a much higher priority than was previously justified. In the process of creating the index maps, an important ancillary benefit was realized – by viewing the size and shape of the bounding box of each map, and by reading the titles and perusing the bibliographic information, previously undetected errors were found and corrected. Although a careful review of each record is time-consuming, it clearly improves the quality of the database.

In order to facilitate a more efficient and routine production of index maps, the data structure for the NGMDB Map Catalog will be revised to incorporate the new fields defined in this study. In cooperation with the agencies that provide new bibliographic records to the Map Catalog, all new entries will include information for these fields. Also, a more automated process of index-map creation will be developed; this process will include the error-checking procedures used here.

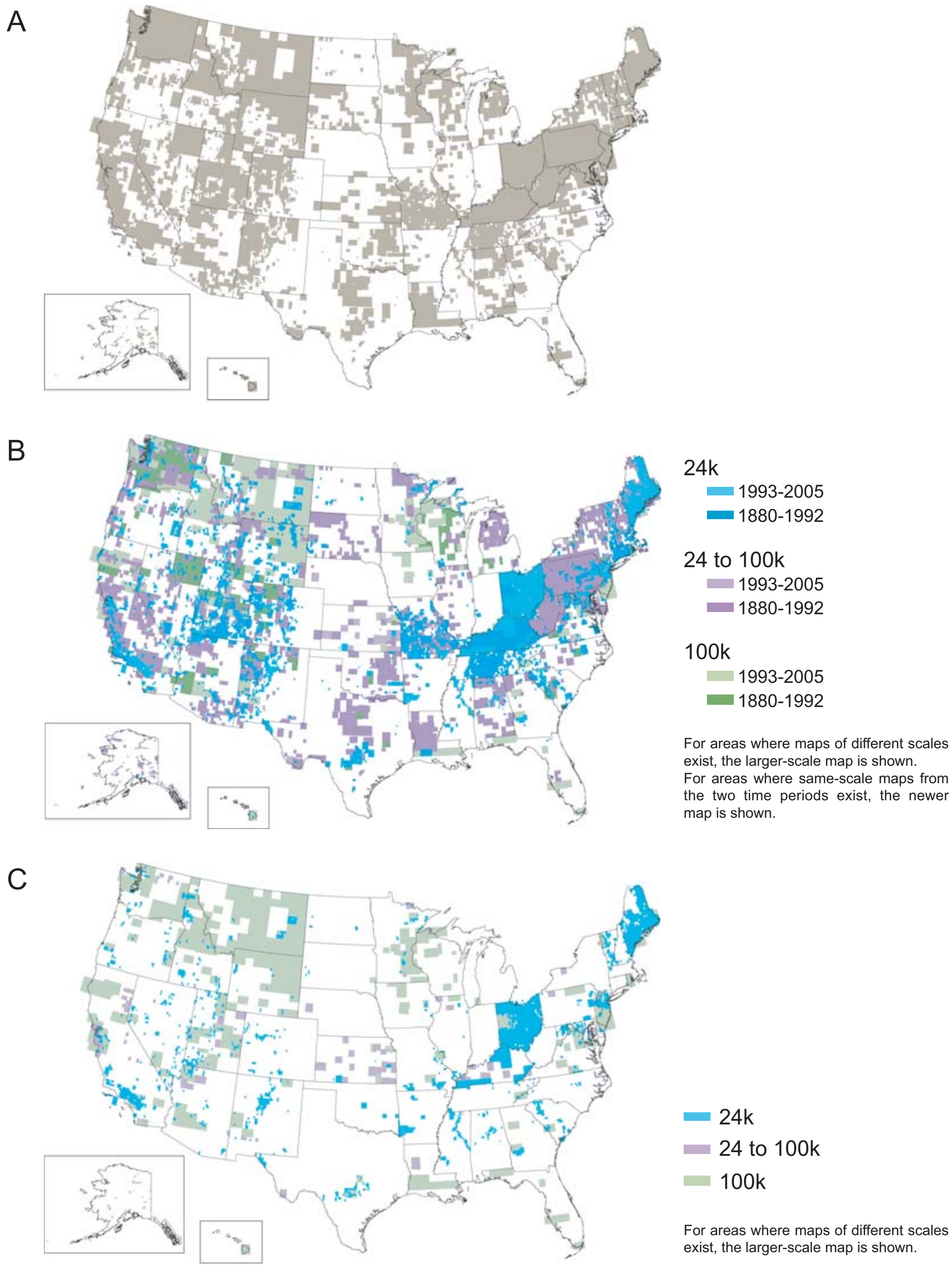


Figure 5. Index maps provided to the NCGMP and AASG, in partial fulfillment of the OMB request for performance metrics.

A) geologic map coverage, at scales 1:100,000 and more detailed, published from 1880 to 2005.

B) geologic map coverage, differentiated into three scale ranges (1:100,000, 1:24,000, and scales between), and published from 1880-1992 or 1993-2005.

C) geologic map coverage, at scales 1:100,000 and more detailed, published from 1993-2005.

Because these index maps show the bounding box for each geologic map there is, for some maps, an overestimation of the area that was actually mapped. In Step 4, it was noted that certain irregularly-shaped geologic maps were included in the index maps. Also, a quadrangle map may cover parts of two states, but the geology may have been mapped for only one state. Therefore, in the future these index maps should more precisely show map boundaries, for example by intersecting certain geologic maps with shoreline and state boundary GIS files, and by representing county and park maps by their true boundaries rather than by bounding box.

I conclude with a somewhat tongue-in-cheek comment. In this study I struggled to infer, from each publication title, certain basic characteristics for each map. My struggle begs the question – do our users, especially the non-scientists, select the most appropriate products from the (sometimes extensive) list of publications shown by a Map Catalog search or an agency's publication list? I suggest that our products might be more readily used if their titles were more succinct and standardized in format and terminology. At this time, I don't have specific recommendations, but as an author I recognize that my own titles could benefit by suggestions for clarity!

REFERENCES

Brabb, E.E., 1989, Geologic map of Santa Cruz County, California: U.S. Geological Survey Miscellaneous Investigations Series Map I-1905, scale 1:62,500, available at http://ngmdb.usgs.gov/Prodesc/proddesc_47.htm

Kepferle, R.C., 1973, Geologic map of the Raywick quadrangle, central Kentucky: U.S. Geological Survey Geologic Quadrangle Map GQ-1048, scale 1:24,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_10593.htm.

McCartan, Lucy, Lemon, E.M., and Weems, R.E., 1984, Geologic map of the area between Charleston and Orangeburg, South Carolina: U.S. Geological Survey Miscellaneous Investigations Series Map I-1472, scale 1:250,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_9253.htm.

Powell, J.W., 1888, Methods of geologic cartography in use by the United States Geological Survey, in Congrès Géologique International, Compte Rendu de la 3me Session, Berlin, 1885: A.W. Schade's Buchdruckerei, Berlin, p. 221–240.

Southworth, C.S., Brezinski, D.K., Orndorff, R.K., Chirico, P.G., and Lagueux, K., 2001, Geology of the Chesapeake and Ohio Canal National Historical Park and Potomac River Corridor, District of Columbia, Maryland, West Virginia, and Virginia: U.S. Geological Survey Open-File Report 01-188, scale 1:24,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_42277.htm.